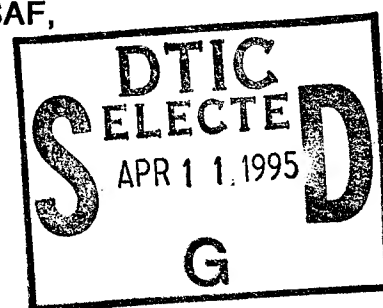RL-TR-94-220
In-House Report
December 1994

OPTICAL CORRELATOR OPTIMIZATIONS AND EXTENSIONS FOR AUTOMATED OPTICAL TARGET RECOGNITION

Wesley E. Foor, Harold G. Andrews II, Capt, USAF, and Mark A. Getbehead

DTIC
SELECTED
APR 1 1 1995
G

Rome Laboratory
Air Force Materiel Command
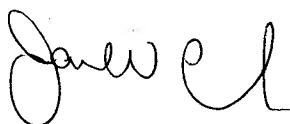**Griffiss Air Force Base, New York**

19950407 153

DTIC QUALITY INSPECTED 5

This report has been reviewed by the Rome Laboratory Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RL-TR-94-220 has been reviewed and is approved for publication.

APPROVED:

JAMES W. CUSACK, Chief
Photonics & Optics Division
Surveillance & Photonics Directorate

FOR THE COMMANDER:

DONALD W. HANSON
Director of Surveillance & Photonics

If your address has changed or if you wish to be removed from the Rome Laboratory mailing list, or if the addressee is no longer employed by your organization, please notify RL (OCPA) Griffiss AFB NY 13441. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE December 1994	3. REPORT TYPE AND DATES COVERED In-House Oct 90 - Oct 93

4. TITLE AND SUBTITLE
OPTICAL CORRELATOR OPTIMIZATIONS AND EXTENSIONS FOR
AUTOMATED OPTICAL TARGET RECOGNITION

5. FUNDING NUMBERS
PE - 62702F
PR - 4600
TA - P1
WU - 09

6. AUTHOR(S)
Wesley E. Foor, Harold G. Andrews II, Capt, USAF and
Mark A. Getbehead

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Rome Laboratory (OCPA)
25 Electronic Pky
Griffiss AFB NY 13441-4515

8. PERFORMING ORGANIZATION REPORT NUMBER

RL-TR-94-220

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Rome Laboratory (OCPA)
25 Electronic Pky
Griffiss AFB NY 13441-4515

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Rome Laboratory Project Engineer: Wesley E. Foor/OCPA (315) 330-2944

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This report presents a detailed overview of the research performed by the analog optical signal processing group at the Photonics Center of Rome Laboratory in the area of automated optical target recognition. This research encompasses phase-only filter optical correlator filter development, optical image remapping, optical image segmentation, and optical neural networks. While the focus of the work is directed toward military applications we also present potential commercial uses for many of these systems.

14. SUBJECT TERMS
optical correlation, optical image remapping, optical filter
design, fractal dimension

15. NUMBER OF PAGES
44

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT U/L

Contents

i

List of Figures

1

INTRODUCTION

Many manufacturing, technical, and military organizations are looking towards machine vision to improve the performance capabilities of automated machines for a large variety of tasks. Real-time pattern recognition is critical to certain applications and as more sophisticated machines and sensors are developed higher processing rates for larger amounts of data are required. Due to the large processing times associated with sequential processors these systems will be most useful for solving real-time problems if they are implemented in parallel hardware. Optical information processing systems offer parallel computation and high-density non-interfering interconnections.

Correlation is a basic operation required by many machine vision and pattern recognition systems. In electronic systems correlations are typically performed using specialized hardware since the algorithms involved are computationally intensive to calculate. Data can be processed at high rates by utilizing the massive parallelism and high bandwidth offered by optics technologies. Current optical correlator systems can perform greater than 1000 correlations per second for 256×256 pixel images with commercially available devices. Near-term advances in optical device development are expected to greatly increase the data processing rates of future systems.

The Photonics Center at Rome Laboratory currently uses an optical correlator based on binary phase-only filters (BPOFs) to develop and evaluate optical pattern recognition algorithms for military applications such as Hostile Target Identification (HTI). Potential civilian applications of this system are finger-print identification for building security, handwritten character recognition for the postal service, string matching for content-addressable-memories, and object detection and recognition for guiding unmanned robots or vehicles. Medical applications include human cell classification and genome searches in human DNA sequences to locate possible genetic defects. The use of an optical processor in these applications presents the opportunity for real-time data analysis.

In general, correlation is not invariant to either affine object distortions such as scaling or rotation or to nonrepeatable distortions of an object's image by atmospheric conditions, diurnal temperature variations, shadows, etc. This presents a major problem for any template-based recognition system where the image of an object to be recognized may change in its appearance. To obtain distortion invariance in a correlator system for real-world problems (e.g. HTI and machine vision) a large library of templates containing distorted versions of the object is typically employed. Large template libraries increase both memory storage requirements and search times. Both of these

1

characteristics are undesirable in a real-time, compact system for use in military applications. We present various approaches to minimizing these problems in an optical correlator system.

This report presents an overview of the work performed by the analog optical signal processing group at the Photonics Center of Rome Laboratory from October 1990 until September 1994. During this time the group has published numerous technical papers and reports in the area of optical signal processing [1]-[24]. In addition to the authors the other members and contributors to the group (both past and present) are Denise Blanchard, Dr. George Brost, Sandy Halby, Capt. Christopher Keefer, and Jackie Smith. It should be noted that in addition to our in-house technical staff, Dr. Samuel Kozaitis from the Florida Institute of Technology (FIT) has played a key role in our optical correlator research and development. This report serves to provide an integrated viewpoint to our work and formally records the work that has not been included in previous Rome Laboratory technical reports. Section 2 of the report outlines our work in the area of phase-only filter optical correlation and describes a photorefractive image correlator. We discuss reduced-resolution optical correlator architectures with the intent of developing faster, less expensive, and more compact systems. Our work on reduced-resolution filter correlators has been expanded upon by P. C. Miller[25] and we summarize his work. In Section 3 we present an alternative optical preprocessor system that makes use of a coordinate transformation to provide a rotation and scale invariant image space. In Section 4 we outline an optical system that performs image segmentation based on a fractal dimension estimation algorithm. In Section 5 we present an optical neural network classifier and finally we present conclusions and possible future directions for our work.

2

OPTICAL CORRELATION

The most common optical correlator is based on a 4f (four focal length) system architecture as shown in Figure 2.1. A typical correlator is comprised of a spatial light modulator (SLM) in both the input and filter planes, two Fourier transforming lenses, and a CCD camera. We begin by describing architectural optimizations of the 4f correlator to reduce the system size, weight, and cost. We then discuss optical filter design. While many algorithms exist for computing optical spatial filters our work has focused on the binary phase-only filter (BPOF) which can be implemented on binary-valued SLMs. BPOFs are useful for recognizing fixed objects in stationary backgrounds and exhibit large, narrow correlation peaks as well as effective multi-class discrimination. They can work well in the presence of background clutter or when an object is partially obscured. In addition, they have been used to identify and track multiple objects. BPOFs have provided suitable solutions when objects have a repeatable signature. If an object varies in a limited or known manner, more complex filters can be used.

2.1 REDUCED-RESOLUTION OPTICAL CORRELATOR

We have proposed a reduced-resolution filter SLM for use in an optical correlator to reduce SLM addressing times and memory storage requirements [16]. The reduction in filter resolution is shown to primarily affect the impulse response by addition of copies of the filter image. This problem can be minimized by placing a diffraction grating in contact with the filter SLM while retaining the advantages of speed and memory storage. Another advantage of this system is the opportunity to use shorter focal length lenses to produce a shorter, more compact correlator. Finally the use of a lower resolution SLMs can greatly reduce the cost of a correlator.

Reducing the amount of data contained in a spatial filter has been previously considered by others. One approach considers the resolution limitations of SLMs to make optical correlators more practical[26]. Another approach reduces the passband of a POF to yield a maximized signal-to-noise ratio. We use a filter SLM that exploits the full bandwidth but has a lower resolution than the input SLM. Our approach results in a constant reduction in the amount of data required to describe a filter.

3
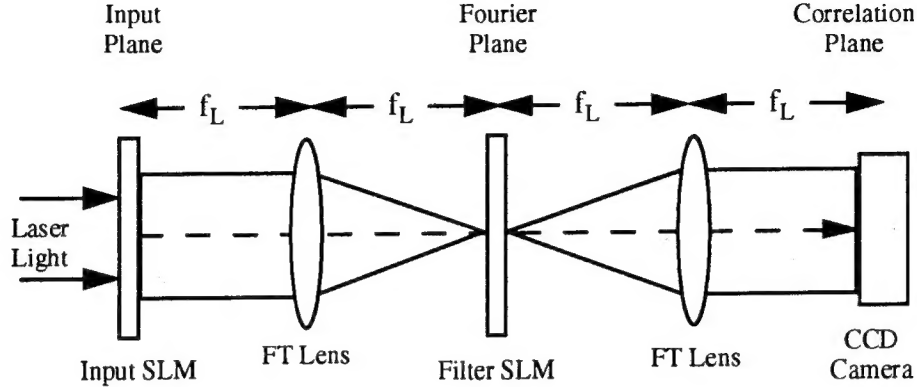
Figure 2.1: Schematic of a 4f optical correlator.

Although a reduced resolution filter correlator may have advantages, its use results in a loss of information. The lower resolution filter can affect the size and detail of the image used to make a filter. Knowing the distortion incurred by using a lower resolution device allows its effects to be minimized.

In this project, we determined the distortion introduced by using a filter that is reduced in resolution by a factor of M along a linear dimension in an optical correlator. In this way, the amount of data used to describe the filter is decreased by a factor of M^2. We provide general expressions for the impulse response, signal-to-noise ratio (SNR), and signal-to-clutter ratio (S/C) for filters reduced in resolution. We also provided guidelines in terms of the size and location of a filter object for minimizing the effect of reducing the resolution of a filter. Furthermore, we show by example the effect of reducing the resolution of a filter in both autocorrelation experiments and cross-correlation experiments containing competing objects. Using these results, the effect of reducing the resolution of other filters can be predicted for other input scenes. The physical implications in terms of the effect on the focal length of lenses in the correlator are also determined.

The effect of using a filter of an integer factor lower in resolution of the input plane in a correlator can be determined using digital signal processing techniques. We create a reduced-resolution filter in the discrete Fourier domain by downsampling a discrete $N \times N$ matched filter by a factor of M where $\{M = 2^i; i = 1, 2, \ldots, log_2 N\}$. More generally, the process is described as

$$Y\left(\frac{k}{M}, \frac{l}{M}\right) = R[X(k,l)],$$

where k and l are integer values and $-N/2 < k, l < N/2$, $X(k,l)$ is the original discrete matched filter; $Y(k/M, l/M)$ is the filter reduced in resolution and is referred to as the optical half-resolution filter; and R is a resolution operator that defines how the samples of the optical reduced-resolution filter are to be calculated. For the above equation, $Y(k/M, l/M)$ will be defined only when k/M and l/M have integer values.

We made use of an FFT algorithm to analyze the effect of the reduced-resolution filters. However, the FFT algorithm requires the same number of samples for both the input and output. Therefore, to perform simulations using the FFT, the optical reduced-resolution filter was extended by interpolation in order to have the same number of samples as a filter at full resolution.

The reduced-resolution filter uses the full bandwidth of the input image; one pixel of a reduced-resolution filter corresponds to M^2 pixels of the full-resolution filter. The optical reduced-resolution filter will result in a constant data reduction by a factor of M^2. It can be represented digitally by $N \times N$ samples with blocks of $M \times M$ samples having the same value. The interpolated $N \times N$ version of the optical reduced-resolution filter is referred to as a reduced-resolution filter.

As a result of the downsampling technique used to obtain the reduced-resolution filters multiple copies of the filter image are in the impulse response of the reduced-resolution filter. Therefore, the copies will also correlate with the input image and multiple correlation responses can appear in the correlation plane. Because the impulse response of the filter is known, it can be used to determine the performance of a correlator as a function of resolution for specific input objects. In addition, we provide guidelines so that results for specific cases can be predicted.

We consider the case in which an object used to make the filter is centered and a translated object in the input plane produces a correlation peak with a full-resolution filter at a distance $k_0, l_0, (k_0, l_0 \geq 0)$ from the center of the correlation plane. When using a reduced-resolution filter, secondary correlation responses will occur at locations

$$\sum_{i=0}^{M-1} \sum_{h=0}^{M-1} \left(k_0 - \frac{iN}{M}, l_0 - \frac{hN}{M} \right).$$

The magnitude of the secondary correlation peaks depend on both the size and location of the object used to make the filter. The impulse response of a reduced-resolution filter will be least attenuated near the center, therefore an input image will correlate more with the central portion of an image used to make a filter. Therefore, we restricted our discussion to the case when an object used to make a filter is centered. Furthermore, the correlation response due to copies of an object will tend to be minimized for relatively small objects since the copies appeared centered at minima in the impulse response.

The secondary correlation peaks can be significant if the object used to make the filter is large enough. An object should be restricted to a $N/M \times N/M$ pixel region to avoid aliasing in the correlation plane. The amount of aliasing that is tolerable is dependent on the specific images used in the correlator. Therefore, the correlation response is dependent on the size of the object used to make the filter. Although the object used to make the filter must be restricted in size, the object to be identified can be located anywhere in the $N \times N$ region of an input image.

The decreased resolution of the filter SLM will change the focal lengths of the lenses in the correlator. Shortened correlators have been previously demonstrated; however, they have not shown a decrease in focal length[27]. The change in focal length depends on the number of pixels in the filter and the specific SLMs being used. The focal length of the Fourier transform lens on the correlator is given as [27]

$$f = \frac{N_2 d_1 d_2}{\lambda},$$

where N_2 is the number of pixels along a side of the filter SLM; d_1 and d_2 are the center-to-center spacing of the pixels in the input and filter plane SLMs, respectively; and λ is the wavelength of the light being used. Note that the focal length of the system is proportional to N_2 if other factors remain constant.

Our experiments demonstrate that under certain conditions, reducing the resolution of a BPOF can be used in an optical correlator with similar performance as compared to full-resolution filter performance. The significance is that the amount of data needed for a reduced-resolution filter is M^2 less than a full-resolution filter where M is the factor that the filter was reduced by in a linear dimension. Further details of the reduced-resolution filter construction, performance measures, and results can be found in references [16, 23].

Recently, our work on reduced-resolution filter design has been extended using an optimal design technique for making the filters more robust to the resolution reduction [25]. In accordance with our results, Miller reports that the reference object (target) size in terms of pixels is the most important parameter with respect to the filter reduction. Miller's optimal design technique provides greater reductions in filter resolution by factors of 4 to 16 over our downsampling techniques. The optimal design technique involves high-pass filtering the reference object imagery as a preprocessing step in the filter construction. His findings indicate that for a given target size, filters constructed with high-pass imagery were marginally more robust to greater amounts of resolution reduction than filters constructed with unprocessed target images.

2.2 PYRAMIDAL PROCESSING

A limiting factor in the application of optical correlators is that the number of pixels in currently available SLMs is often not large enough for military applications. One approach is to process large images at lower resolutions beginning with the lowest resolution. Pyramidal processing is a multiresolution processing technique in which an image is processed multiple times, each time at a different resolution. In pyramidal processing an image is represented by a series of lower resolution versions of itself up to and including the original image.

A common way to generate an image pyramid is to filter an image with Gaussian functions of increasing standard deviations and downsample the image by increasing amounts. First, an original $2N \times 2N$ image is low-pass filtered to eliminate spatial frequencies greater in magnitude than $p/2$ where the spatial frequencies of the image extend from $-p$ to $+p$. The image is then downsampled by a factor of 2 to yield a $2N - 1 \times 2N - 1$ image. To generate other levels of the pyramid, the $2N \times 2N$ image is filtered to eliminate spatial frequencies greater in magnitude than p/L, then downsampled by a factor of L to yield a $2N - L \times 2N - L$ image. If the original image $f(m,n)$ is referred to as level 0, then $f_L(m,n)$ represents different versions of the original image at level L, where $0 < m, n < 2N - L - 1$, and L is the level of the pyramid $\{L = 0, 1, 2, \ldots, N\}$.

Generally, low-pass filtering produces grey-leveled values at some pixels. Since grey levels cannot be implemented on binary SLMs, we considered an alternative way to generate a pyramid structure. Morphological processing is used here instead of Gaussian low-pass filtering to create a pyramid structure. In a morphological pyramid, an opening operation can be treated as a low-pass (size) filter that performs smoothing of the contour of an object. In addition, the output of the opening

6

operation is binary so that the resulting opened and downsampled image can be displayed on an SLM.

Images of $LN \times LN$ pixels can be processed in parallel with an optical correlator using SLMs of $N \times N$ resolution by employing a pyramidal processing technique. The use of morphological operators to generate a pyramid image representation allows the technique to be implemented with binary SLMs and appears to give similar results to that of low-pass filtering. Our results show that the SNR and discrimination decrease with increasing L and that this technique is limited by the decrease in SNR. The SNR is primarily affected by the scaling factor due to the downsampling.

Our experiments show that we can use a correlator with 128×128 pixel resolution to successfully identify an object in a 512×512 image. In our experiments the use of morphological operators allow binary correlators to achieve SNR values within a factor of one-half that of the maximum result obtained when an ideal low-pass filter had been used. The discrimination abilities are also comparable to the ideal results. Due to the relatively small number of pixels in currently available SLMs, this technique is useful for the processing of large images in spite of the added complexity of the technique. More details of our research in multi-resolution processing for optical correlators can be found in references [3, 22].

2.3   DISTORTION INVARIANT OPTICAL FILTER DESIGN

A major difficulty encountered when using a BPOF in an optical correlator is its sensitivity to changes in the object's appearance. Images of an object can vary significantly depending on aspect angle, lighting, atmospheric effects, and a host of other variables. In addition, object boundaries may be poorly defined and buried in the background. Identifying an object that has a nonrepeatable signature is one of the key technical challenges of automatic object recognition.

BPOFs are useful for recognizing fixed objects in stationary backgrounds. BPOFs exhibit large and narrow correlation peaks and provide effective multi-class discrimination[28, 29]. They can work well in the presence of background clutter or when an object is partially obscured[30]. BPOFs have provided suitable solutions when objects have a repeatable signature. If the image of an object varies in a limited or known manner, more complex filters, such as the synthetic discriminant function (SDF), can be used.

2.3.1   Synthetic discriminant function (SDF) optical filters

SDF optical filters are created by using linear combinations of the input reference images during filter construction. Using this method it is possible to overcome much of the correlator's sensitivity to object distortions. If the distortions are predictable or repeatable (e.g. scale and rotation distortion), then it is possible to create a filter or small set of filters to recognize each of the distorted views of the object. Typically we train the SDF filter in an iterative mode adjusting the relative weight of each object in the linear combination in order to obtain a constant correlation peak height for each reference input. By iteratively training the SDF filter, system noise and imperfections can be learned, effectively performing an on–line calibration of the system. We have

7

investigated SDF filters for threshold invariance[17], scale and rotation invariance, and invariances to other distortions[21].

This section contains a brief discussion of how the SDF filters are made. A more detailed description of the approach we followed has been previously reported [31]. Assuming that a specified correlation response is produced for a set of training images, the need for displaying different filters is reduced or eliminated. A conventional SDF is a weighted combination of images that can be described as

$$s(x,y) = \sum_n a_n t_n(x,y),$$

where t_n are centered training images and a_n are weight coefficients. SDF synthesis techniques may be used to determine the weight coefficients [32, 31]. The complex conjugate of the Fourier transform of $s(x,y)$ is the matched filter

$$S(u,v) = \mathcal{F}[s(x,y)]^*,$$

where $\mathcal{F}$ is the Fourier transform operator.

Converting the SDF-matched filter to a BPOF may result in a severe loss of information. Recently, an improved version of an SDF, called a filter SDF (fSDF), has been introduced that includes the function modulation characteristics of the device onto which the filter is mapped in the synthesis equations [31]. For fSDF-BPOFs, the coefficients an can be iteratively trained based on the formula

$$a_n^{i+1} = a_n^i + \beta \left[c_n - c_0 \left(\frac{m_n^i}{m_0^i} \right) \right], \tag{2.1}$$

where i is the iteration number, β is a damping constant, and m_n^i is the modulus of the peak correlation response of image $t_n(x,y)$ with a filter made with the coefficient vector $\mathbf{a}^i$. In the experiments described in this report, the initial solution vector was taken to be the desired correlation response vector, $\mathbf{a}^0 = \mathbf{c} = 1$. The initial fSDF, $s(x,y)$, was then found and cross-correlated with each training image. The values of the correlation heights were then placed into Eq. 2.1, and an updated $\mathbf{a}$ was found. A new $s(x,y)$ was found and the procedure is repeated. The modulation characteristics of an a BPOF SLM was included by calculating intermediate cross-correlation in Eq. 2.1 with BPOFs.

2.3.2 SDF filters applied to thresholded imagery

An optical correlator that uses binary SLMs requires the conversion of sensor imagery to binary imagery. The conversion process is highly vulnerable to noise and variations of the object and background. Therefore, an object can appear differently after the image is converted to a binary image due to environmental or other conditions. The reliability of the conversion process is critical for object recognition because the binary image contains shape features of the object. In real–world imagery, the global shape of an object is frequently too perturbed to generate a reliable, specific version of the object. The binary result is often a version of the object that changes in an unknown or nonrepeatable way.

An automatic method is needed for detecting objects with BPOFs using imagery from infrared (IR) sensors. By using digital image processing techniques, images can be confined by simulation

to vary in a limited but unknown manner. SDF-BPOFs are then used to identify objects. It is sometimes difficult to evaluate the performance of a distortion-invariant method because of the lack of a suitable performance measure. Therefore, to help evaluate the real potential of this approach, imagery is presented from actual sensors that were not from the original training set.

We present a method for identifying objects in infrared imagery using SDF-BPOFs. The method is suitable for applications that involve objects with a nonrepeatable signature. Rotation and scale invariance should be achieved by storing a bank of filters that are rotated and scaled versions of filters developed here.

IR imagery (8-14 μm) of ground scenes from actual sensors were used to evaluate the proposed method. Images were digitized with 128×128 pixels with 8 bits/pixel. Because the application was to binary SLMs, the imagery had to be thresholded.

Thresholding was performed by choosing a single threshold value for the entire image. Threshold values can be chosen several different ways. They can be based on the noise statistics of an image, a histogram of the image, or a fixed value chosen near the middle of the available pixel values. When the object and background are within an image are obvious, a threshold value can be easily chosen, and different methods usually give similar results. Other techniques can be used if the imagery is more complex. In either case, a thresholding method should be automatic in that it should perform similarly with a variety of imagery under various lighting and atmospheric conditions.

In the imagery we examined, the background and object were easily separated; however, edges between them were not well-defined. We used digital image processing techniques to implement a thresholding method. This isodata technique examines peak values in the histogram of an image. A threshold value was chosen between peaks that were associated with the object and the background so that the object could be segmented from the background. If noise or atmospheric distortion was present, the peaks of the histogram would change their position or shape, but they will usually be identified. Choosing a threshold value between peaks of a histogram often results in an image that is similar to the silhouette of the object. As variables such as lighting, noise, and atmospheric effects within an image change, resulting thresholded images will remain similar but will often be different in an unknown or nonrepeatable way.

Different threshold values of an image containing an object may produce different binary images. An IR photograph with a $0.3m$ ground resolution was digitized so that each pixel corresponds to a $1.4m \times 1.4m$ area and is shown in Figure 2.2. Figures 2.3a-2.3d show versions of Figure 2.2 thresholded at different values; Figures 2.3a and 2.3d are thresholded at values 84 and 108, respectively, where black and white pixels have values of 0 and 255. Images that were thresholded with values between 84 and 108 produced a different image for each value; values outside this range produced severely distorted images and were not considered. There is just less than a 1000-pixel difference in the number of white pixels in Figures 2.3a and 2.3d.

The IR image shown in Figure 2.2 was used to produce training images for several fSDF filters. Threshold values between 84 and 108 were used to generate the training images. SDFs were made that had three, five, and seven training images. Many different combinations of images were used to generate fSDFs using the algorithm in Eq. 2.1. An fSDF-BPOF was produced for every possible training set, and one filter was chosen for a training set of a fixed number. To choose a representative fSDF for a given number of training images, thresholded versions of Figure 2.2 at values between
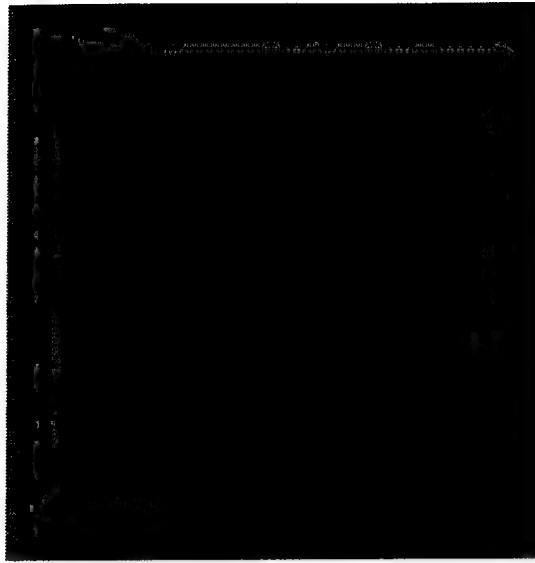
Figure 2.2: Infrared image from which training sets were derived.

Filter	Threshold values (Range 0–255)
SDF3	92, 96, 100
SDF5	88, 92, 96, 100, 104
SDF7	84, 88, 92, 96, 100, 104, 108

Table 2.1: Threshold values for image of Figure 2.2 used to produce training images for SDF filters.

84 and 108 were correlated with each SDF filter and the correlation response examined. The fSDF filter that produced the most consistent correlation heights for the range of thresholded images was chosen for further analysis. Using this procedure, one fSDF filter was chosen for each of the three, five, and seven training images. The threshold values of Figure 2.2 used to produce the training images used in our experiments are shown in Table 2.1. An increase in the number of training images added images with threshold values outside the range of the original training set. Therefore, the addition of training images extended the distortion range of the filter.

Results show that the distortion range of BPOFs was increased to useful amounts for automatic object recognition using the preprocessing described here and the proper choice of training set for an fSDF-BPOF. Using an image from an actual sensor not in the training set, the level for the decision criterion for the best BPOF increased between 6.8% to 13% depending on the fSDF filter used. Had the BPOF been made from a sample image not threshold at the optimum level, the decision criteria for the poorest performing BPOF increased between 22.3% to 28.5% .

Because actual sensor data varied in an unknown way, the choice of the optimum training set for the fSDF-BPOFs was not straightforward. However, linearly independent images for SDF filters can be created from a set of training images using an orthogonalization process. Decreasing the
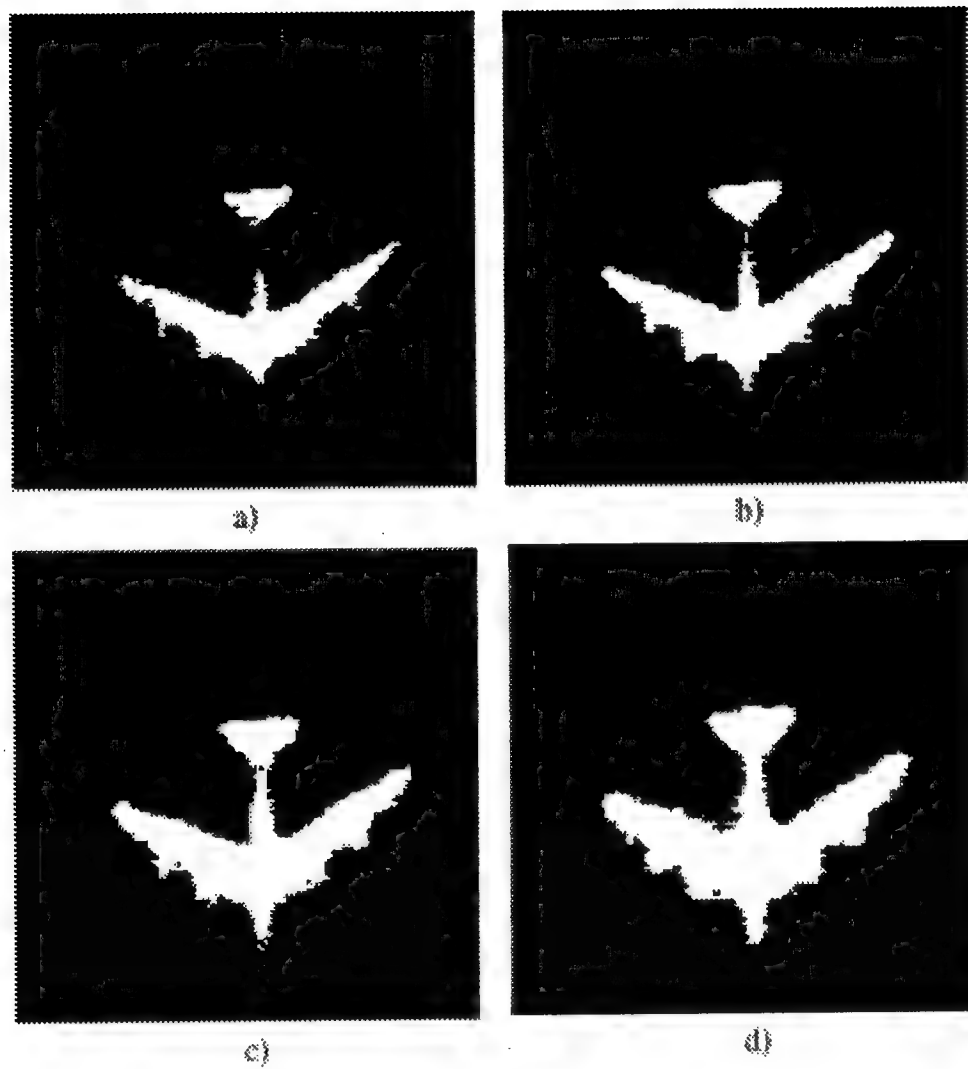
Figure 2.3: Thresholded versions of Figure 2.2. Thresholded at gray-level values: a) 84, b) 92, c) 100, and d) 108.

sensitivity of a BPOF to the different ways that an input can vary is essential to automatic object recognition using BPOFs. Further details of this work can be found in reference [17].

2.3.3 Feature-based optical filters

Using an optical correlator, we experimentally evaluated a BPOF designed to recognize objects not in the training set used to design the filter. Such a filter is essential for recognizing objects from actual sensors. We used an approach is that is as descriptive as a BPOF yet robust to object and background variations of an unknown or nonrepeatable type. We generated our filter by comparing the values of spatial frequencies of a training set. Our filter was easily calculated and offered potentially superior performance to other correlation filters.

We investigated the use of a BPOF that was calculated from an objects features. In this way, we attempted to make a BPOF more robust to unknown variations than other designs. We have previously shown by computer simulation that our filter offered superior performance to SDF filters [20] for our problem; here, we provide experimental results for a version of our filter. We considered a one-class problem; our results were generated using a training set from only one class of objects. To help evaluate the potential of our approach, we used imagery from actual sensors that were not from the original training set.

In actual sensor imagery, the global shape of an object is frequently too distorted to generate a specific version of the object. Therefore, an input object may not correlate well with a filter even though the input and filter are from the same class [17, 33]. In contrast to SDF filter formulation, we developed a filter whose values were determined by features of the objects in a training set. We attempted to find a filter that represented the critical characteristics of a class of objects so that objects outside the training set but in the same class could be identified. Therefore, we examined features that were invariant with respect to the training set.

Generally, the cross-correlation between two images is maximized when the mean squared error (MSE) between the images is minimized. The correlation operation measures the similarity of images; therefore, images will correlate well if their Euclidean distance in signal space is small. In signal space, an image is represented as a point and each axis may represent a spatial frequency. A point along an axis represents the value of that spatial frequency. The region in signal space that represents images that correlate well with a given image is generally a multidimensional sphere centered on the given image. The radius of the sphere is determined by a threshold in the correlation plane where any correlation response above the threshold is considered to be a match with the given image. As the threshold decreases, the radius of the sphere will increase.

We attempted to form n training images into a cluster in signal space by retaining only spatial frequencies with a small spread of values. We examined the Discrete Fourier Transforms (DFTs) of the training images at each spatial frequency. The DFT of the kth training image was represented by $S_k[u, v]$ where u and v are discrete spatial frequencies. The values of the spatial frequencies were examined across the entire training set in terms of their similarity. We considered the distance between their values in the complex plane as a measure of their similarity. The smaller the distance, the more similar the values.

In conclusion the feature-based filter offered a range of performance. In the case where none of

the pixels were set to zero in the filter, the fSDF and feature-based filter offered similar performance. The feature-based filter was slightly more consistent and had broader correlation peaks for objects within the training set than the fSDF filter. Neither filter appeared to be useful for recognizing objects outside the training set.

As pixels of the filter were set to zero in the feature-based filter, the correlation peaks within the training set became more consistent even though their average height decreased. As the number of pixels set to zero increased, the correlation heights became more consistent but broader. When images of the same class as the training set but not in the training set were used as inputs, the feature-based filter was potentially useful. Our experiments involved five training images. The use of more training images suggests that more possibilities are available in trading off between consistency and broadness of the correlation results. In this way, the feature-based filter can be made robust to recognized object outside the training set.

2.3.4 Ternary POFs

We developed ternary phase-only filters that identified objects outside a training set in the presence of unknown or nonrepeatable distortions. In our experiments, our statistical filters recognized objects within the same class and in the presence of noise better than another popular binary distortion-invariant filter design.

In contrast to previous attempts at distortion-invariant BPOF formulation, we developed a filter whose values were determined by features of the training set. We attempted to find a filter that represented the critical characteristics of an object so that objects outside the training set could be identified. Because correlation filters are derived from the Fourier transform of an object, we examined feature extraction in the Fourier domain. We considered the BPOFs of input images as a set of features to recognize objects and used a statistical approach to examine features that were invariant with respect to the training set. We retained those Fourier features that were invariant among a training set, and set to zero those that varied using a technique similar to factor analysis to design a ternary filter.

The principle components method, which is related to factor analysis, has been used to design correlation filters. In contrast, we examined an ensemble of BPOFs to select spatial frequencies to recognize images outside of our ensemble.

Our statistical filters offered a range of performance depending on a parameter p. In the case $p = 1$ (none of the pixels set to zero), the fSDF filter offered slightly better performance. However, as p increased the correlation peaks within the training set became more consistent and their normalized average correlation height increased. When images in the same class as, but not in the training set were used as inputs, our statistical filters had a higher normalized average correlation height than the fSDF filter. Furthermore, in most cases, our statistical filters produced more consistent and higher normalized average correlation heights than the fSDF filter in the presence of noise. Therefore, in our experiments, our statistical filters recognized objects within the same class and in the presence of noise better than the fSDF filters.

Our statistically designed filters were more easily calculated than an fSDF filter. Our filters required on the order of calculating N FFTs. In contrast, the fSDF approach requires a cross-

correlation between every training image and filter every iteration. This requires on the order of Ng FFT calculations where g is the number of iterations and has been generally set to ten [31]. Therefore, the time and number of operations required to calculate our statistical filters were about an order of magnitude less than for the fSDF filter.

Because our statistically designed filter recognizes an object based on its features, other objects may be identified if they have similar features. Therefore, it may be important to include discrimination ability into our statistically designed filters. Further details of our work can be found in reference [10].

2.4   PHOTOREFRACTIVE IMAGE CORRELATOR

This section briefly describes a correlator fabricated by Accuwave of Santa Monica, CA under direct support from Rome Laboratory. Details of this work are given in Rome Laboratory Technical Report RL-TR-94-154 [2]. The holograms were recorded at Accuwave and the system was tested in-house at the Photonics Center. The optical image correlator uses orthogonal, wavelength multiplexed Fourier transform holograms recorded in a photorefractive crystal. Cross-correlation and auto-correlation measurements were obtained using randomly selected test images against a set of reference images stored in the orthogonal data storage volume hologram. More than 40 holograms were written in the 645–$651nm$ wavelength range with $> 2\%$ diffraction efficiency and 1.5 Å wavelength separation. A low power tunable external cavity semiconductor laser was used for hologram readout, demonstrating the portability of the approach. The input image translation tolerances on the correlation output and the effect of using partial images during readout were also investigated.

This type of correlator uses a fixed set of reference filters recorded in the volume hologram. There is an upper limit on the number of filters that can be stored in the crystal and that may restrict this system architecture to applications that have a well defined feature space to correlate against. The primary advantage of this system is that the entire template database is 'static' and no computer interface is required to change the correlation templates. This differs from the typical correlation systems which are limited in speed by the filter SLM/computer interface. The speed limitations of this correlator are influenced by the detector integration time and the rate at which the laser can be tuned.

3

OPTICAL LOG-POLAR COORDINATE TRANSFORM PREPROCESSOR

Log-polar coordinate transforms are a well established technique for highlighting specific scale and rotation properties of an object of interest[34]. In the human vision system the eye-to-brain mapping is a log-polar mapping process[35]. The log-polar coordinate transform presented in this report provides a feature space where Cartesian angular position is remapped to the x-axis and the radius is remapped to the y-axis. When the remapped image of the object of interest is used in a correlator, object variations in rotation and scale are represented as linear shifts on their respective axis of the correlation output plane. These linear shifts of the correlation peak provide information about the size and rotation of the object with respect to the reference filter. In machine vision applications and image processing, estimation of the orientation and size of an object are important tasks. It has also been shown how the log-polar transform can simplify the direct estimation of the time to impact for autonomous vehicles[35].

A number of authors propose using log-polar remapping to overcome the problem of recognizing objects which vary in rotation and scale [34, 36]. The optical image remapper we describe can be integrated with a correlator or neural network for the purpose of determining the scale and rotation of a particular object. An advantage of the optical image remapper versus electronic implementations is the ability to perform the log-polar transformation of an image in parallel, theoretically faster than possible with a digital electronic processor. Potential applications of this optical remapper include machine vision for identifying known objects in various orientations and target recognition of objects on the battlefield from high altitude surveillance platforms.

Casasent and Psaltis proposed a feature space that is invariant to scale and rotation changes in order to reduce the number of correlation filters required for a recognition task[34]. The key to their approach is to perform a coordinate transformation in order that the scale and rotation of a given object is represented in a new feature space or coordinate axis system. This feature space would be mapped in rectangular coordinates with $\ln(r)$ mapped along one axis, where r is radial position in Cartesian coordinates, and the angular position θ mapped along the orthogonal axis.
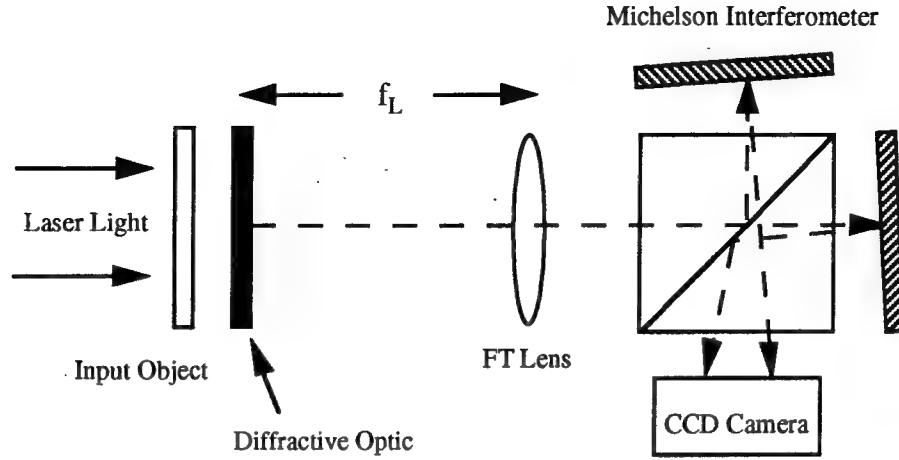
Figure 3.1: Schematic of optical system to perform log-polar coordinate transform.

The new coordinate axes can be represented as

$$\ln(r) = \ln \sqrt{(x^2 + y^2)}$$

and

$$\theta = \arctan(y/x).$$

The major limitations in using the coordinate transformation in machine vision or pattern recognition is that it only works well for a single object and the object must be centered in the Cartesian coordinate system before remapping into the log-polar coordinate system. There are many existing digital recognition schemes that detect blobs (possible objects) within a scene and then create a region of interest (ROI) about the centroid of the object. This blob detection technique can also be performed using an optical correlator. In general this recognition scheme would solve the single object per scene limitation and the centering problem. Another method of centering the object within a scene is to perform a Fourier transform (FT) of the scene and record only the magnitude of the spectrum [34]. This eliminates the linear phase terms in the spatial frequency domain which correspond to the shifted positions of the object within the spatial domain.

The coordinate transformation is performed by the combination of a computer generated hologram (CGH) and a Fourier transforming (FT) lens as shown in Figure 3.1. The CGH has a phase transmission ϕ given by

$$\phi(x, y) = \left(\frac{-2\pi x_0}{\lambda f_L} \right) \left[x \ln \sqrt{(x^2 + y^2)} - y \arctan(y/x) - x \right],$$

where x_0 is a constant of the same units as x and y, λ is the laser wavelength, and f_L is focal length of the FT lens. A plot of the phase transmission is shown in Figure 3.2. The CGH used in the experimental set-up is a binary phase level device with 1000 points in each direction, x and y, in a feature space of a $10mm^2$. The constant x_0 is set to $1mm$. Our system design is intended for use with a HeNe laser ($\lambda = 632.8nm$) and the FT lens design focal length is $200mm$. A picture of the actual transparent CGH device is shown in Figure 3.3.
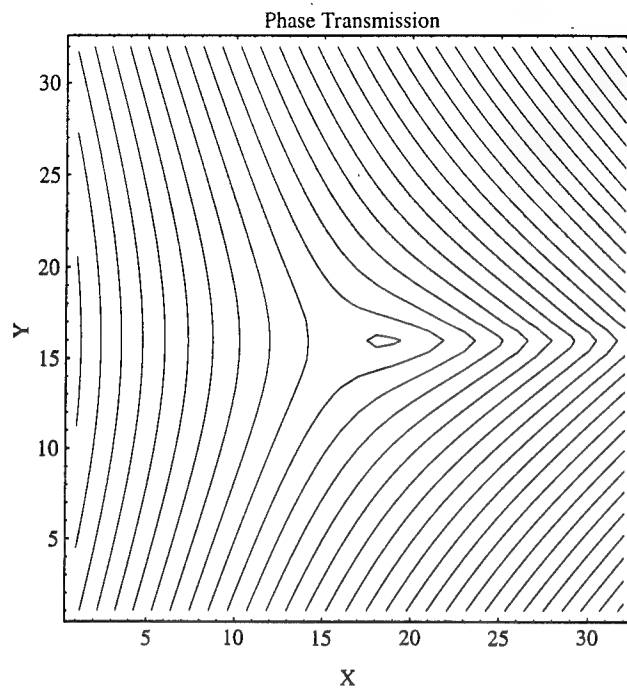
16

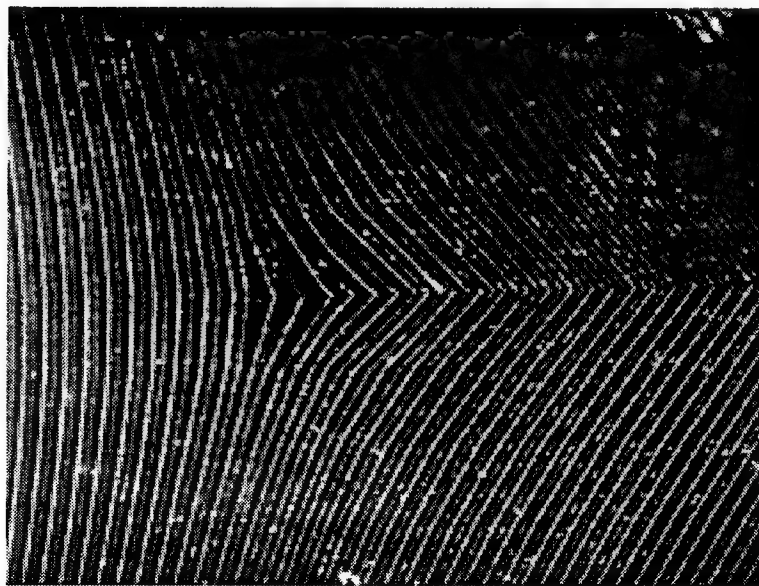Figure 3.2: Plot of the phase transmission for a 32x32 point array.

Figure 3.3: Picture at 100X magnification of the central portion of the CGH used to perform the log-polar coordinate transform.

x

y

ln(r)

θ

a)

b)
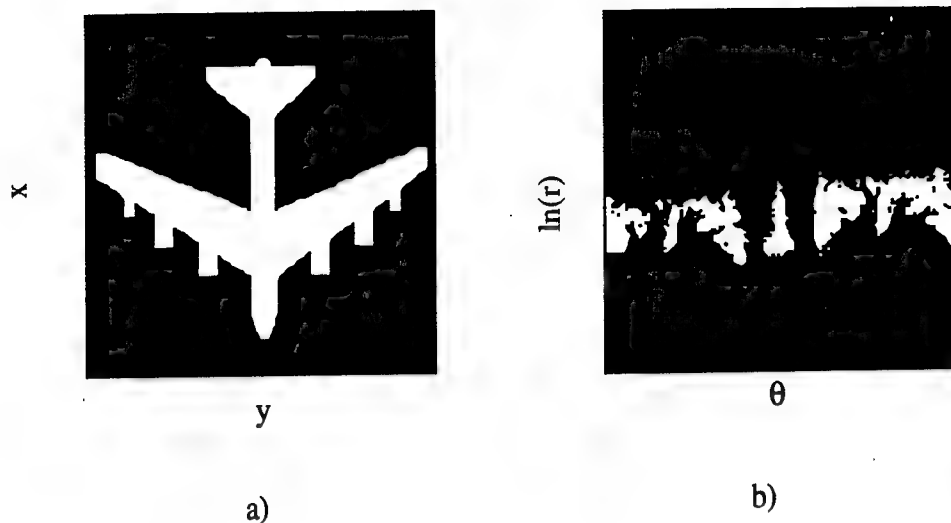
Figure 3.4: a) Sample input image and b) the output of the optical log-polar preprocessor.

Again referring to Figure 3.1 we note that the CGH is placed in close proximity to the input SLM in order to minimize diffraction effects. A Michelson interferometer is used to obtain the superimposed images of the coordinate transformation in the Fourier plane of the FT lens. Figure 3.4 shows a sample input image and the corresponding output of the log-polar optical preprocessor. The Michelson interferometer produces two identical patterns on the polar axis by displacing the interferometric path with respect to the horizontal axis. There is a slight region of interference between the two paths where overlap occurs. This overlap interference causes a fringe pattern which degrades the FT image slightly. A CCD camera collects the intensity of the coordinate transform output at a distance f_L from the lens.

An optical correlator employing this system as a front-end preprocessor, as shown in Figure 3.5, can then use a single log-polar reference filter to compare to the unknown input image. In this system the location of a positive correlation peak can be used to calculate the identified object's scale and rotation[13]. Utilizing this preprocessing system can greatly reduce the number of filters required to perform object recognition. Further details of the system and its performance can be found in reference [6].

Figure 3.5: Optical system incorporating log-polar preprocessor and a 4f correlator.

4

OPTICAL FRACTAL DIMENSION ESTIMATION
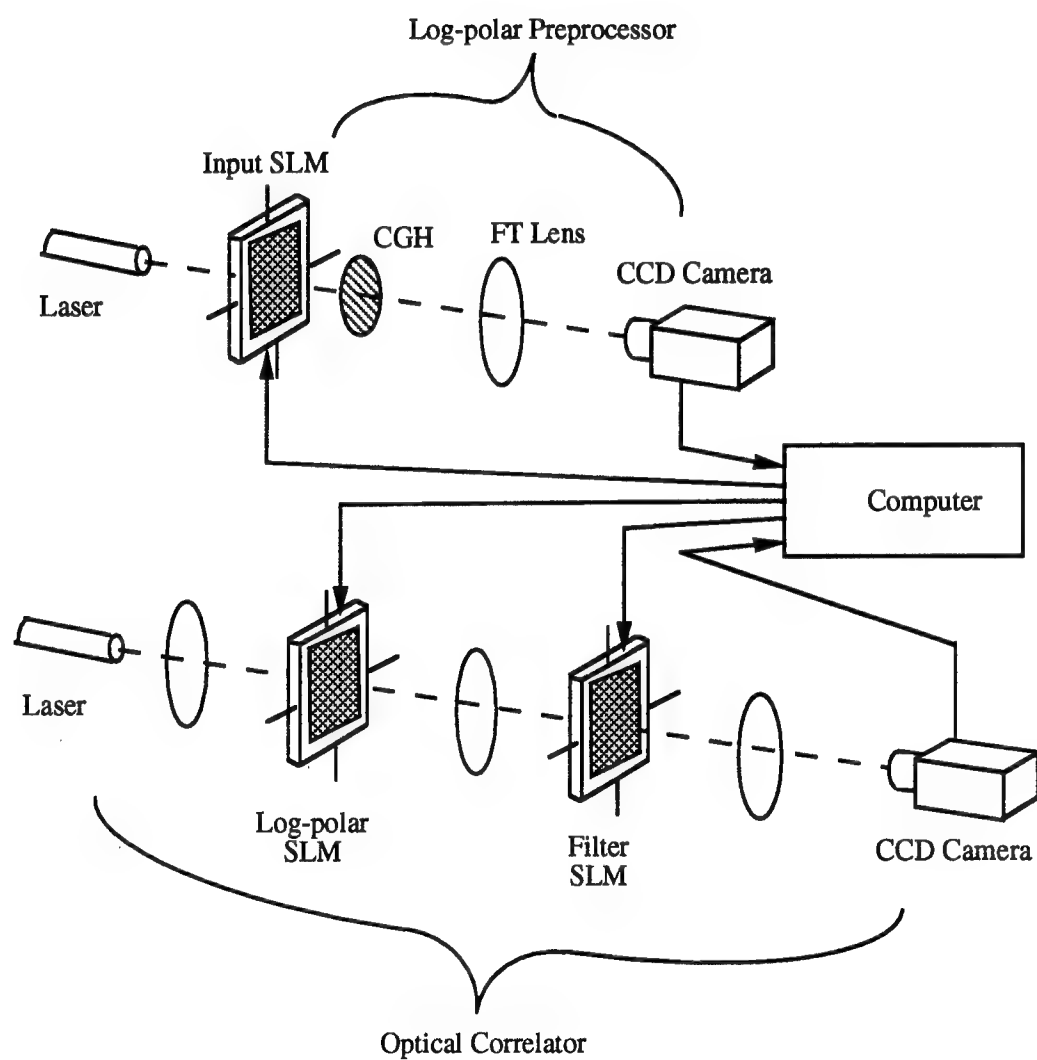
This research investigated the use of fractal dimension measure to segment spatially disjoint regions of interest from simulated fractal clutter or background [5, 12, 18]. The underlying assumption is that a given region of interest in a real-world image has a different fractal dimension than its background. We investigated virtually illuminated, digitally simulated fractal surfaces with known fractal dimensions. The backgrounds we considered had various degrees of texture roughness. We constructed an optically based image segmentation system to perform the otherwise computationally intensive Fourier transform of the image to be segmented. We compared the performance of this system to an all digital approach. Though useful for such things as aerial and space based reconnaissance, there are many other applications that could also benefit from the techniques described here. For instance, when applied to machine vision applications, these techniques could help reduce the time required to locate some tool against spatially disjoint clutter. They could also prove useful to applications involving robotic navigation of guidance for hazardous material cleanup. In both cases additional processing will allow the machine to make decisions based on information from a few regions of interest. These techniques could also possibly prove useful as a preprocessor of imagery generated by medical scanners. The rationale is that a growth may have a different fractal dimension than the surrounding tissue.

Previous theoretical and experimental work [37, 12] established a relationship between the topological features of a fractal surface, the surface's illuminated image, and its power spectrum. From these relationships, we estimate a fractal dimension measure from an optical Fourier transform and digital post-processing its power spectrum. From these results, certain inferences can be drawn concerning the location of regions of interest. Namely, the techniques discussed here can quickly spot features having different fractal dimensions from the surrounding clutter.

This investigation compared the ability of an all digital technique to a hybrid optical-digital technique for estimating the fractal dimension of the computer generated imagery. The digital method took a fast Fourier transform of the illuminated image, and used that to calculate the image's fractal dimension. The optical-digital technique did essentially the same thing, though the Fourier transform was taken optically. A Fourier lens, a 256×256 Semetex Magneto-Optic SLM

and a CCD camera at the Fourier plane composed the optical system. Once we had the Fourier transform, digital post-processing calculated the fractal dimension of the original illuminated image. This digital post-processing was identical in both the digital system and the hybrid optical-digital system.

4.1 SURFACE GENERATION AND ILLUMINATION

We used the spectral synthesis method to generate fractal surfaces [38]. First, we generated two dimensional random Fourier components $G_{i,j}$ with a mean amplitude of zero and a standard deviation

$$\sigma^2 = (i^2 + j^2)^{-(H+1)/2},$$

on the $(i, j)th$ random discrete Fourier component, where H is related to the desired fractal dimension D [18] of the surface by $D = 2 - H$. A computer was used to perform an inverse Fourier transform to generate a fractal surface g where

$$g(x, y) = \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} G_{k,l} \exp\left[2\pi i(kx + ly)\right].$$

Surface g was illuminated using a pure Lambertian model where the intensity at a particular location $I(x, y)$ is given by

$$I(x, y) = \cos(p_{x,y}),$$

where $p_{x,y}$ is the angle between the normal of g at (x_0, y_0) and the direction to the infinitely distant point source illuminant. The normal $\mathbf{N_0}$ at (x_0, y_0) is

$$\mathbf{N_0} = g_x(x_0, y_0)\mathbf{i} + g_y(x_0, y_0)\mathbf{j} - \mathbf{k}.$$

We calculated the power spectrum, $P_H(f, t)$ of I by summing the squares of amplitudes within particular frequency rings. We then band-pass filtered the power spectrum, plotted it on a log-log graph, and fitted it to a line. A linear relationship exists between the slope of the line $-m$, and the fractal dimension D of the original illuminated image [37] where

$$D = 3 - \frac{m}{2}.$$

Twelve surfaces were created using the spectral synthesis method described above. Illuminating each surface from a variety of angles required knowledge of the normal to the surface at each of the $65,536$ points composing the surface. We derived the normal of $g(x_0, y_0)$ from the partial derivatives as described above. To get these partial derivatives we used discrete Fourier transforms to numerically approximate these partial derivatives. This approach was very compute intensive, and limited the number of Fourier components we could use.

Six of these twelve surfaces used 16×16 Fourier components and the other six used 32×32 components. For each set of surfaces, we varied the H parameter from 0.0 to 1.0 in increments of 0.2. A computer virtually illuminated each of these twelve surfaces from six angles, and generated simulated imagery as viewed from directly above. The six angles of illumination varied from $0°$ to
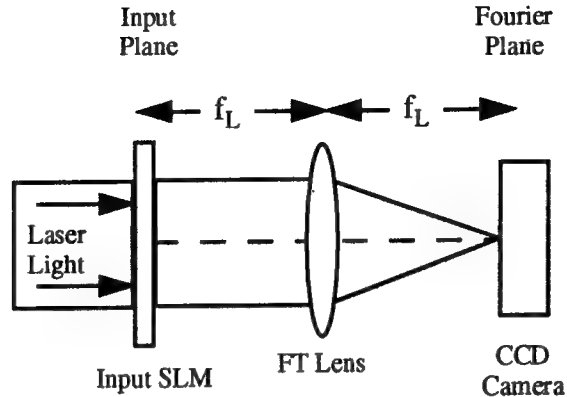
Figure 4.1: Schematic of optical fractal dimension estimation system.

90° relative to the viewing angle. We stored the 72 resulting images in 256×256 BMP format gray scale files.

We should note here that the surfaces images are not self shadowing. To establish certain baseline characteristics of the algorithm and the optical system's performance, we decided to remain consistent with the notions established in previous literature on this subject [37]. Additionally, the surfaces considered followed the properties of a pure Lambertian illumination model for the same reason.

4.2 OPTICAL VS. DIGITAL FRACTAL DIMENSION ESTIMATION

An optical system like that shown in Figure 4.1 performed a Fourier transform of each of the 72 images. This setup could take the Fourier transform only of binarized images since Semetex 256×256 SLM used in our experiment is a binary device. As such, we thresholded the grayscale images at their average intensity level before placing them onto the SLM. The Fourier transform of the image on the SLM was imaged onto the CCD camera. A frame grabber card then captured this image into a personal computer. We then clipped and placed the image from the camera into a binary file for image processing.

The digital technique used the fast Fourier transform (FFT) routines in the Image Pro Plus software package running on a 33MHz 80486DX computer. With Image Pro we calculated the FFT of all the illuminated surface images and stored the amplitude information in BMP binaries. The phase information was discarded. Each 256×256 Fourier transform required approximately five seconds to compute. Figure 4.2 shows a typical image from its surface contour map, to its illuminated image, to the image of its FFT. An example of the images taken from the optical system was not easily ported into this report.

Additional processing calculated the power spectrum of each of the 144 Fourier images and
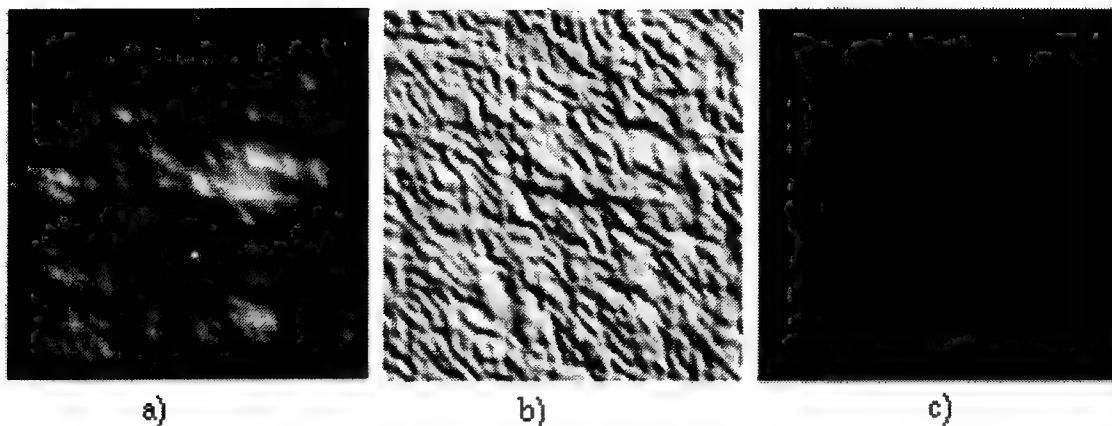
Figure 4.2: Example of image generation and processing. a) Contour map, b) illuminated image, and c) Fourier Transform of illuminated image.

H/t	0°	18°	36°	54°	72°	90°
0.0	2.043	2.005	1.944	1.906	1.905	1.940
0.2	2.123	2.094	2.024	1.973	1.964	2.002
0.4	2.198	2.172	2.102	2.049	2.038	2.070
0.6	2.236	2.231	2.211	2.188	2.183	2.192
0.8	2.011	2.044	2.114	2.140	2.128	2.109
1.0	1.549	1.606	1.739	1.864	1.870	1.793

Table 4.1: Digital fractal dimension results D for 16×16 Fourier components with parameter H and illumination angle t.

saved the data to ASCII files. To reduce the effects of noise with the optical system (arising mainly from the pixelation of the SLM), we blocked the sections of the optical Fourier transform extending both horizontally and vertically from the DC. A computer digitally bandpass filtered all of the Fourier transforms and graphed the resulting power spectra on a log-log plot.

The slope of the line fit to the data in the log-log plots is $-m$. Tables 4.1 through 4.4 show the values for D in each of the 144 Fourier transforms. The H value refers to the parameter for generating the fractal surface, while t refers to the angle of the illumination. Except extreme cases in illumination angle or the parameter H, the digitally computed values cluster closely to each other for a given fractal dimension. The range of fractal dimension results for a particular value of H does not intrude upon the range calculated for another value of H, though it does occasionally occur.

We now consider occluded fractal surfaces illuminated from a variety of angles. A geometric shape (e.g. square) can be placed over part of the illuminated image to see how this changes the fractal dimension measure D from the non-occluded imagery (see Figure 4.3). This was done with

H/t	0°	18°	36°	54°	72°	90°
0.0	1.799	1.819	1.781	1.745	1.744	1.766
0.2	1.91	1.92	1.876	1.832	1.828	1.860
0.4	2.043	2.074	2.033	1.986	1.974	2.002
0.6	2.244	2.307	2.286	2.246	2.230	2.260
0.8	2.261	2.432	2.510	2.522	2.518	2.509
1.0	1.751	2.072	2.311	2.434	2.442	2.351

Table 4.2: Digital fractal dimension results D for 32×32 Fourier components with parameter H and illumination angle t.

H/t	0°	18°	36°	54°	72°	90°
0.0	1.707	1.765	1.652	1.613	1.704	1.776
0.2	1.745	1.713	1.732	1.685	1.865	1.824
0.4	1.735	1.740	1.700	1.809	1.828	1.844
0.6	1.853	1.791	1.790	1.760	1.805	1.836
0.8	1.845	1.806	1.876	1.823	1.872	1.900
1.0	1.765	1.989	2.013	1.960	1.996	1.965

Table 4.3: Optical fractal dimension results D for 16×16 Fourier components with parameter H and illumination angle t.

H/t	0°	18°	36°	54°	72°	90°
0.0	1.517	1.803	1.804	1.715	1.651	1.784
0.2	1.503	1.804	1.803	1.581	1.667	1.868
0.4	1.448	1.658	1.494	1.514	1.709	1.837
0.6	1.473	1.459	1.527	1.602	1.784	1.828
0.8	1.402	1.718	1.611	1.752	1.817	1.847
1.0	1.759	1.885	1.814	1.833	1.823	1.873

Table 4.4: Optical fractal dimension results D for 32×32 Fourier components with parameter H and illumination angle t.
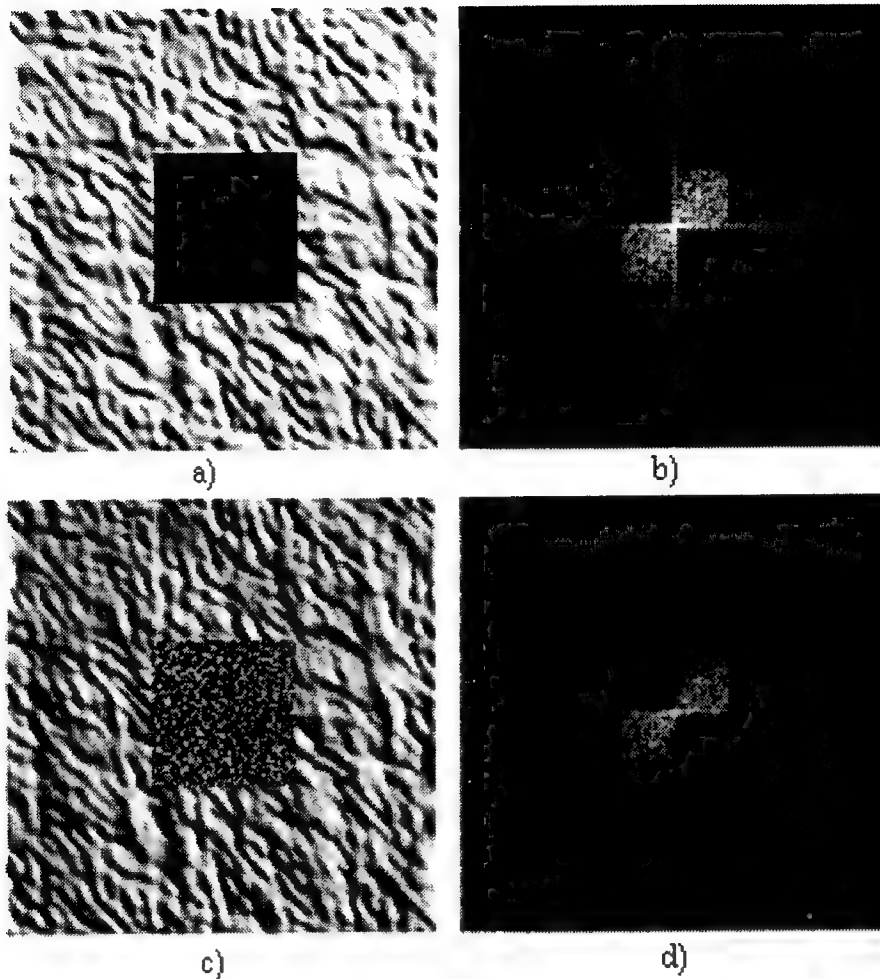
24

Figure 4.3: a) Uniform pulse image and b) its Fourier Transform, c) random pulse image and d) its Fourier Transform.

both a uniformly shaded square covering the middle of the selected images, and with a square region filled with random 8-bit values. (We call them a uniform pulse and a random pulse respectively). Tables 4.5 through 4.8 show values for D when we employed the two techniques on the two sets of surfaces.

In each case, we handled the images identically to those in the non-occluded case, and reduced the data in the same fashion. Due to the large amount of energy in the higher frequencies of the random pulse, the slope increased, decreasing the value of m (flattening things out a bit). Similarly, there was a great deal of spectral energy along the axes, characteristic of sharp edges, and a large value at the DC.

The digital approach seems well suited to differentiate between the two pulse types and the unpulsed data in both the 16×16 and 32×32 Fourier component tests. The difference between the maximum and minimum values is rarely greater than 0.1 save for values of $H = 1.0$. This implies

H/t	18°	36°	54°	72°
0.4 (Random)	1.164	1.212	1.248	1.249
0.4 (Uniform)	2.015	1.994	1.979	1.976
0.6 (Random)	1.097	1.124	1.150	1.145
0.6 (Uniform)	1.992	2.018	2.033	2.051

Table 4.5: Digital fractal dimension results D for 16×16 Fourier components with parameter H, illumination angle t, and uniform or random pulse as indicated.

H/t	18°	36°	54°	72°
0.4 (Random)	1.330	1.418	1.462	1.467
0.4 (Uniform)	1.973	1.969	1.942	1.940
0.6 (Random)	1.233	1.322	1.370	1.370
0.6 (Uniform)	2.071	2.106	2.113	2.126

Table 4.6: Digital fractal dimension results D for 32×32 Fourier components with parameter H, illumination angle t, and uniform or random pulse as indicated.

that a deviation greater than 0.1 may show a potential region of interest, and may warrant further investigation by either human or electronic processing. Most of the uniform pulse images were at least 0.1 from all of the unpulsed images with that H value. All of the random pulses were even further away.

Turning our attention to the optical setup, we see that there is a bit of a reduction in the ability to discriminate the random pulse and unpulsed data. However, usually there is still little overlap between the two. Here, we discriminate the uniform pulse much more easily than in the all digital process. The difference in performance characteristics may have been the result of noise in the SLM. Upon viewing the output from the SLM, there were several lines of light, parts of which should have been turned off. Also, the light passing throughout the SLM at the region containing the random pulse did not appear distributed properly. This may have contributed to the poor performance. Pixelation was not as much of a factor as it could have been. As noted earlier, we digitally blocked the axes when calculating the power spectrum. This should have reduced, if not

H/t	18°	36°	54°	72°
0.4 (Random)	1.688	1.739	1.616	1.590
0.4 (Uniform)	2.183	2.016	1.958	1.960
0.6 (Random)	1.779	1.692	1.681	1.691
0.6 (Uniform)	2.224	2.062	2.037	1.992

Table 4.7: Optical fractal dimension results D for 16×16 Fourier components with parameter H, illumination angle t, and uniform or random pulse as indicated.

H/t	18°	36°	54°	72°
0.4 (Random)	1.790	1.867	1.832	1.810
0.4 (Uniform)	2.185	2.084	2.066	2.036
0.6 (Random)	2.005	1.870	1.785	1.735
0.6 (Uniform)	2.071	2.106	2.113	2.126

Table 4.8: Optical fractal dimension results D for 32×32 Fourier components with parameter H, illumination angle t, and uniform or random pulse as indicated.

eliminated the effects of pixelation. However, it also removed the spectral energy we expected to see in the uniform pulse.

4.3 PERFORMANCE COMPARISON

Each approach described has strengths and weaknesses. The primary advantage of the optical system is that the potential speed is far greater than that offered by any of the reasonably priced digital alternatives. Semetex claims a 50fps frame rate on its 256i device. Thus, 250 Fourier transforms can be performed optically in the time it takes to calculate one FFT on the digital platform used here. This assumes that computer hardware controlling the optical system can retrieve imagery at a minimum of 50fps.

There were several disadvantages when using the optical system as well. The optical system seemed more prone to noise, and the Semetex requires a great deal of fine tuning to get the image displayed properly on the device. Incorrect switching of entire rows and columns of pixels, SLM pixelation, and nonuniform illumination of the SLM all combined to produce noise at the detector array. Optical aberrations and imperfect alignment generated crosstalk, further degrading discrimination ability. Additionally, the Semetex often requires more than one write to the array to eliminate large horizontal bands of light from passing through the device. We failed to match Semetex's 50fps frame rate on the 256i.

The advantages of the digital system involve the ability to reduce the noise levels of the process. Based on the results in the tables below, for a given H value, the fractal dimension measure D has less variation in the digital system than the optical. FFTs have been sufficiently debugged, and computer performance has become increasingly cost-effective that digital FFTs are offering a serious challenge to the speed benefit derived from optical image processing. This is especially the case when considering the time and resources required to write information to the SLM and read that data back from the CCD array. Also, the digital system does not require throwing away as much information as does the optical system. Digital systems use eight bit gray scale data, while optical systems require us to eliminate seven of those eight bits.

The main disadvantage offered by the digital approach compared with the optical system is speed. Optical systems have the potential to outperform digital systems since they compute Fourier transforms at the speed of light. However, severe bottlenecks exist when writing an image to the

SLM and reading its Fourier Transform from the camera. Improvements in optical device and analog to digital conversion technology may overcome these bottlenecks.

Based on the performance of the techniques discussed, it is possible to segment square pulses from fractal backgrounds based on the fractal dimension measure. This implies that there is some merit in considering how this approach deals with more sophisticated shapes occluding portions of more realistic scenery. The next logical step in this line of investigation is to look at scanning across high resolution imagery to detect areas where abrupt changes in the fractal dimension occur.

Though the capability to view even small (256×256) optical images in anything approaching a real-time fashion is expensive with the off-the-shelf technology, this technology has applications to other areas in which that capability is not much of a consideration. These techniques could be employed to highlight regions within images taken by various pieces of medical scanning equipment or to automate the process of searching for regions of interest within aerial or space based reconnaissance imagery. Applications requiring real-time image processing may benefit from these techniques when the optical device technologies mature sufficiently.

5

OPTICAL NEURAL NETWORK CLASSIFIER

Current work focuses on using the adaptive nature of neural networks to compensate for optical imperfections and noise in classifier systems as well as creating a more compact feature representation. The inherent parallel processing capabilities of opto-electronic processors and the relatively simple computational requirements of artificial neural networks make optics a good candidate for hardware implementation of neural networks. Neural networks require large numbers of interconnections for which optics can facilitate high-density, non-interfering parallel connections in free space. Optical systems suffer from the problems of noise and optical imperfections for which on-line learning can provide additional system robustness.

The radial basis function (RBF) neural network is an adaptive system that learns on-line and has been successfully used in many multi-dimensional classification applications including radar signal classification[39, 40, 41], 3D object recognition[42], speech recognition[43], and handwritten character recognition[44, 1]. It has been reported that while having equal or better performance than back-propagation neural networks on classification tasks, the training times for RBF networks are much shorter[45].

A hybrid opto-electronic implementation of a RBF neural network as shown in Figure 5.1 has been demonstrated as an adaptive real-time classifier/interpolator [1, 7]. Using a bipolar encoding scheme a true Euclidean distance computation can be performed all optically (see Figure 5.2). The current implementation requires binary inputs and binary node locations. This binary vector representation is well suited to feature-based 3D object recognition, character recognition, and certain other multi-dimensional classification tasks. In this feed-forward network the learning of the weights and the gaussian widths is performed after the distance computation and can be implemented with analog VLSI or DSP technologies. The use of analog VLSI post-processing electronics allows the system to retain the flexibility and accuracy of electronics while letting the optics perform the massively parallel computations and data reduction. This architecture also offers ease of scalability and fully parallel input/output capability for real-time real-world problems. Details of our work in this area can be found in reference [1].
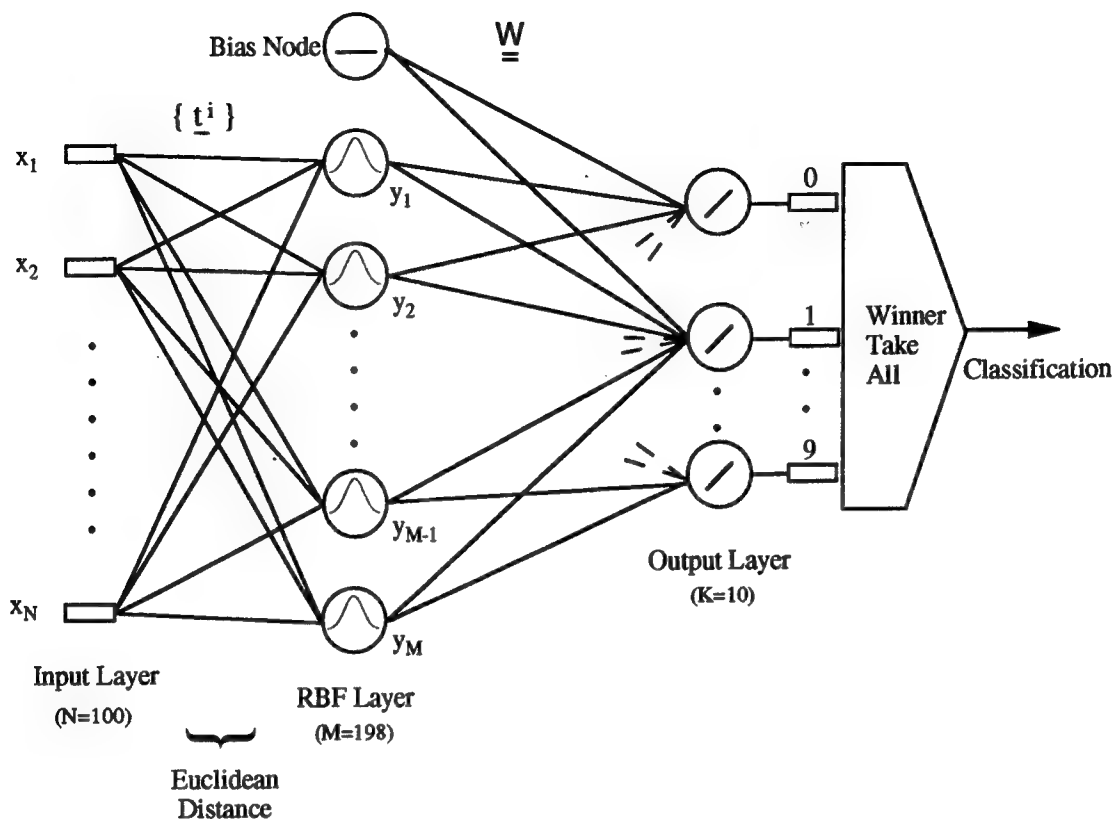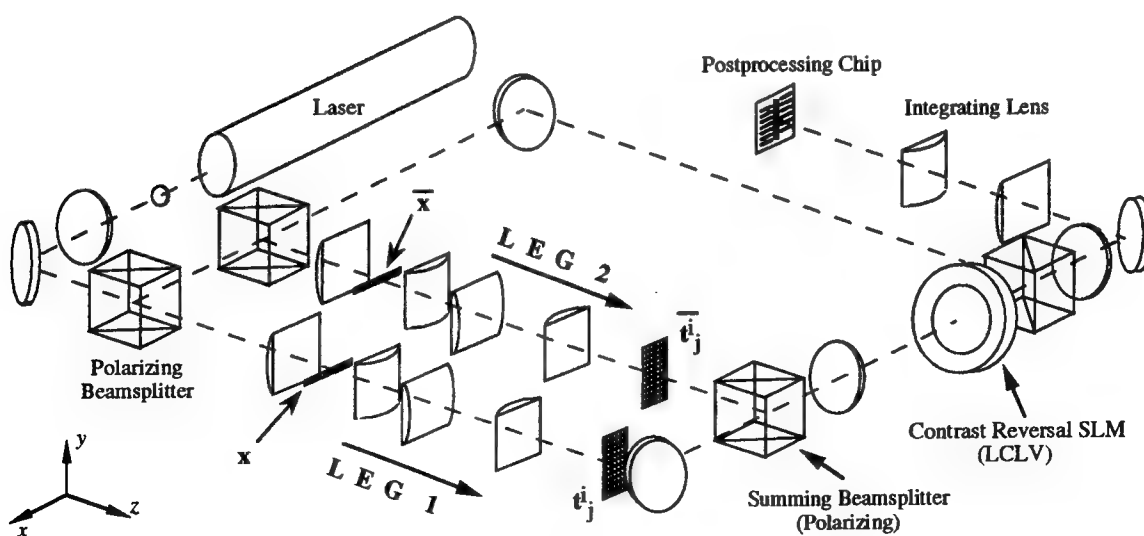
Figure 5.1: Diagram of RBF neural network.

Figure 5.2: Adaptive optical radial basis function neural network.

6

CONCLUSIONS

We have presented two systems for use as real-time signal classifiers. Both the optical correlator and the optical neural network architectures are versatile enough to be applied to many different problems without hardware modifications but by changing only the data representation and reference templates. This implies that a single generic architecture may be fabricated for use in both military and civilian applications. The reduced-resolution optical correlator that we have presented offers the advantages of shorter correlator length, faster SLM addressing, decreased memory requirements, and cost reduction. The optical neural network offers the capabilities of adaptive training and on-line calibration of the optical system as well as faster classification times as compared to correlators.

The problems of object distortion invariance and adaptation to system imperfections have been addressed in order to make the systems more reliable in 'real-world' problems. Due to the large variability in real-world images there exists a need for enormous template libraries for recognizing even a single 3D object. To perform faster recognition in real-world scenarios the trend in correlation-based systems is to incorporate more distortions or simply multiple reference objects into a single template. This is typically done through the use of SDF filters. We have shown that the use of SDF filters can provide selected invariances to distortions typically encountered in Air Force applications. We have also shown that only a few distortions or objects can be placed onto a single filter before the performance becomes intolerable. This leaves us with the need to still perform sequential searches through large template databases when using a correlator. Neural networks offer the capability of encoding the feature space of the entire set of templates into a single system. This enables the neural network system to perform a parallel or 'one-shot' classification of an unknown input.

We have presented two optical pre-processor systems for use with an optical correlator or neural network. The log-polar coordinate transform system offers a scale and rotation invariant feature space which can reduce the number of filters needed for template matching. The optical fractal dimension estimator system which can be useful for quickly identifying smaller regions-of-interest in large scenes.

Our near-term plan is to focus on the development of adaptive optical neural network architectures and algorithms in order to exploit the capabilities of on-line training and system calibration

31

and to perform 'one-shot' classification. Two of the neural network paradigms we are currently developing are the radial basis function [1] and pulse-coupled neural networks [46]. We are also currently investigating the use of wavelet transforms, both optical and electronic implementations, for use as pre- and post-processors for the classifier systems. As the devices used in these optical systems are improved in terms of speed, size, and optical performance the advantages over electronic systems will become more evident for real-time processing.

REFERENCES

[1] W. E. Foor and M. A. Neifeld. Adaptive optical radial basis function neural network for handwritten digit recognition. *Submitted to Appl. Opt.*, September 1994.

[2] K. Sayano and G. A. Rakuljic. Optical Image Correlator Using Orthogonal Data Storage. Rome Laboratory Technical Report 94-154, August 1994.

[3] S. P. Kozaitis and W. E. Foor. Binary optical correlation using pyramidal processing. *Opt. Eng.*, 33(6):1838–1844, June 1994.

[4] W. Foor. Optical neural network for handwritten digit recognition. In *IEEE Dual-Use Technologies and Applications Conference*, volume II, pages 53–58, May 1994.

[5] H. G. Andrews and W. E. Foor. Fractal dimension estimation for machine vision applications. In *IEEE Dual-Use Technologies and Applications Conference*, volume II, pages 59–63, May 1994.

[6] C. W. Keefer, M. A. Getbehead, and W. E. Foor. Log-Polar Optical Coordinate Transformation with Applications for Automatic Pattern Recognition. Rome Laboratory Technical Report 94-22, March 1994.

[7] W. Foor and M. A. Neifeld. Adaptive optical radial basis function neural network for handwritten digit recognition. In *Proc. SPIE*, volume 2240, pages 155–163, April 1994.

[8] W. Foor, M. Getbehead, H. Andrews, C. Keefer, and J. D. Smith. Automated optical target recognition. In *Proc. SPIE*, volume 2216, pages 39–44, April 1994.

[9] S. P. Kozaitis and W. E. Foor. Distortion invariant binary and ternary phase-only filter for distortion invariance using factor analysis. In *Proc. SPIE*, volume 2216, pages 171–180, April 1994.

[10] S. P. Kozaitis, R. H. Cofer, and W. Foor. Statistical design of ternary phase-only filters for distortion invariance. *Opt. Commun.*, 103(1):46–52, 1993.

[11] S. P. Kozaitis. Development of Phase-Only Filters for Sensor Imagery. Rome Laboratory Technical Report 93-248, December 1993.

[12] H. G. Andrews II, M. A. Getbehead, and S. P. Kozaitis. Fractal dimension estimation for optical image segmentation. In *Proc. of SPIE*, volume 2026, pages 361–370, July 1993.

[13] C. W. Keefer, M. A. Getbehead, and W. Foor. Recognizing Objects of Various Rotation and Size with an Optical Image Remapper. In *IEEE Technologies and Applications Conference*, pages 172–176, May 1993.

[14] S. P. Kozaitis and W. E. Foor. Feature-Based Phase-Only Filtering. Rome Laboratory Technical Report 93-30, April 1993.

[15] S. P. Kozaitis, R. H. Cofer, and W. Foor. Design of distortion-invariant correlation filters for sensor imagery using supervised learning. In *Proc. SPIE*, volume 1959, pages 1959–20, 1993.

[16] S. P. Kozaitis and W. Foor. Optical correlation using reduced resolution filters. *Opt. Eng.*, 31(9):1929–1935, 1992.

[17] S. P. Kozaitis and W. Foor. Performance of synthetic discriminant functions for binary phase-only filtering of thresholded imagery. *Opt. Eng.*, 31(4):830–837, 1992.

[18] S. P. Kozaitis, H. G. Andrews, and W. Foor. Optical image analysis using fractal techniques. In *Proc. SPIE*, volume 1790, pages 117–124, 1992.

[19] S. P. Kozaitis and W. Foor. Feature-based correlation filters for object recognition. In *Proc. SPIE*, volume 1790, pages 104–111, 1992.

[20] S. P. Kozaitis and W. Foor. Distortion-invariant correlation using nonlinear feature-based phase-only filters. In *Proc. SPIE*, volume 1772, pages 208–218, 1992.

[21] S. P. Kozaitis, R. Petrilak, and W. Foor. Feature-based correlation filters for distortion invariance. In *Proc. SPIE*, volume 1701, pages 264–273, 1992.

[22] S. P. Kozaitis, Z. Saquib, R. H. Cofer, and W. Foor. Multiresolution template matching using an optical correlator. In *Proc. SPIE*, volume 1702, pages 155–164, 1992.

[23] S. P. Kozaitis, N. Tepedelenlioglu, and W. E. Foor. Analysis and experimental performance of reduced-resolution binary phase-only filters. In *Proc. SPIE*, volume 1564, pages 373–383, July 1991.

[24] D. M. Blanchard. A Performance Model of Thermal Imaging Systems (TISs) Which Include the Human Observer's Response to "State of the Art" Displays. Rome Laboratory Technical Report 91-307, September 1991.

[25] P. C. Miller. Optimum reduced-resolution phase-only filters for extended target recognition. *Opt. Eng.*, 32(11):2890–2898, 1993.

[26] J. Shamir, H. J. Caulfield, and J. Rosen. Pattern recognition using reduced information content filters. *Appl. Opt.*, 26(12):2311–2314, 1987.

[27] J. A. Davis, W. A. Waring, G. W. Bach, R. A. Lilley, and D. M. Cottrell. Compact optical correlator design. *Appl. Opt.*, 28(1):10–11, 1989.

[28] M. A. Flavin and J. L. Horner. Correlation experiments with a binary phase-only filter implemented on a quartz substrate. *Opt. Eng.*, 28(5):470–473, 1989.

[29] J. L. Horner and H. O. Bartlett. Two-bit correlation. *Appl. Opt.*, 24(18):2889–2897, 1985.

[30] K. H. Fielding and J. L. Horner. Clutter effects on optical correlators. In *Proc. SPIE*, volume 1151, pages 130–137, 1989.

[31] D. A. Jared and D. J. Ennis. Inclusion of filter modulation in synthetic discriminant function construction. *Appl. Opt.*, 28(2):232–239, 1989.

[32] D. Casasent. Unified synthetic discriminant function computational formulation. *Appl. Opt.*, 23(10):1620–1627, 1984.

[33] S. P. Kozaitis, S. L. Halby, and W. E. Foor. Experimental performance of a binary phase-only optical correlator using visual and infrared imagery. In *Proc. SPIE*, volume 1296, pages 140–151, April 1990.

[34] D. Casasent and D. Psaltis. Position, rotation, and scale invariant optical correlation. *Appl. Opt.*, 15:1795–1799, July 1976.

[35] M. Tistarelli and G. Sandini. On the advantages of polar and log-polar mapping for direct estimation of time-to-impact from optical flow. *IEEE Trans. Patt. Anal. Machine Intell.*, 15:401–410, April 1993.

[36] D. Asselin and H. H. Arsenault. Rotation invariant pattern recognition using a double coordinate transform. Presented at OSA Annual Meeting, September 1992.

[37] P. Kube and A. Pentland. On the imaging of fractal surfaces. *IEEE Trans. Patt. Anal. Machine Intell.*, 10(5), 1988.

[38] H. O. Peitgen, editor. *The Science of Fractal Images.* Springer-Verlag, Berlin, 1988.

[39] T. O'Donnell, J. Simmers, and H. Southall. An Introduction to Neural Beamforming. In *IEEE Dual-Use Technologies and Applications Conference*, volume I, pages 483–492, May 1994.

[40] J. Simmers and T. O'Donnell. Adaptive RBF Neural Beamforming. In *IEEE Technologies and Applications Conference*, pages 94–98, May 1993.

[41] G. Vrckovnik, C. Carter, and S. Haykin. Radial Basis Function Classification of Impulse Radar Waveforms. In *Proc. of IJCNN*, volume I, pages 45–50, June 1990.

[42] T. Poggio and S. Edelman. A network that learns to recognize three-dimensional objects. *Nature*, 343:263–266, 18 Jan 1990.

[43] S. Renals and R. Rohwer. Phoneme Classification Experiments Using Radial Basis Functions. In *Proc. of IJCNN*, volume I, pages 461–467, Wash. D.C., June 1989.

[44] Yuchun Lee. Handwritten Digit Recognition Using K Nearest-Neighbor, Radial Basis Function, and Backpropagation Neural Networks. *Neural Computation*, 3:440–449, 1991.

[45] J. Moody and C. Darken. Fast Learning in Networks of Locally-tuned Processing Units. *Neural Computation*, 1(2):281–294, 1989.

[46] J. L. Johnson. Pulse-coupled neural nets: translation, rotation, scale, and intensity signal invariance for images. *Appl. Opt.*, 33(26):6239–6253, 1994.

MISSION

OF

ROME LABORATORY

Mission. The mission of Rome Laboratory is to advance the science and technologies of command, control, communications and intelligence and to transition them into systems to meet customer needs. To achieve this, Rome Lab:

    a. Conducts vigorous research, development and test programs in all applicable technologies;

    b. Transitions technology to current and future systems to improve operational capability, readiness, and supportability;

    c. Provides a full range of technical support to Air Force Materiel Command product centers and other Air Force organizations;

    d. Promotes transfer of technology to the private sector;

    e. Maintains leading edge technological expertise in the areas of surveillance, communications, command and control, intelligence, reliability science, electro-magnetic technology, photonics, signal processing, and computational science.

The thrust areas of technical competence include: Surveillance, Communications, Command and Control, Intelligence, Signal Processing, Computer Science and Technology, Electromagnetic Technology, Photonics and Reliability Sciences.